Evidence-based Practice Center Systematic Review Protocol

Project Title: Use of Natriuretic Peptide Measurement in the Management of Heart Failure

I. Background and Objectives for the Systematic Review

Overview

Heart failure (HF) is a major concern for health care systems particularly since it is a chronic condition of older persons and requires substantial resources to manage. HF affects approximately 5.7 million Americans and about 670,000 new cases are diagnosed annually.[1] Based on current population estimates,[2] HF is prevalent in 1.8% of the American population. The estimated total cost for HF in 2010 was $39.2 billion, or 1 to 2 percent of all healthcare expenditures.[1] The aging population along with the need to be efficient with health care dollars requires sound evidence to provide direction for the diagnosis and management of this disease.

One of the advances in HF that has been made is the discovery of the biomarker B-type natriuretic peptide (BNP). This 32 amino acid peptide is released into the bloodstream in response to increased ventricular wall stress, hypertrophy, and volume overload. BNP and N-Terminal pro B-type natriuretic peptide (NT-proBNP) measurement have been included in clinical practice guidelines for HF.[3]

A comprehensive systematic review on BNP and NT-proBNP was completed in 2006 by the McMaster Evidence-based Practice Center (EPC) for the Agency for Healthcare Research and Quality (AHRQ) and included studies up to January 2005.[4] The key questions (KQs) centered on how BNP or NT-proBNP could be used for diagnosis in acute and non-acute settings, its prognostic value for HF and other cardiac events, and if it had value in monitoring of therapy for persons with HF. Currently, use of the BNP and NT-proBNP test varies among health facilities in the United States and other countries. Part of the reason for this discrepancy is how health care is delivered in different places. More importantly, though, the discrepancy exists on account of the wide body of diffuse literature on the subject, especially with regard to the utility of BNP or NT-pro-BNP testing to aid diagnostic and prognostic decision-making in clinical settings. Although the original 2006 review addressed these issues, the number of primary studies in the diagnostic and prognostic fields has increased exponentially since that time. Also, advances in assay types have since rendered some of the publications included in the 2006 review obsolete. For the same reasons, uncertainty surrounds the BNP or NT-pro-BNP test's ability to predict morbidity and mortality in HF patients, as well as whether therapies involving BNP or NT-pro-BNP lead to improved outcomes in HF patients. Indeed, all potential areas of use for BNP and NT-pro-BNP testing require clarification using currently available assays that have been approved by regulatory bodies such as the United States Food and Drug Administration (FDA). We also do not have up-to-date knowledge of whether the biological variation of BNP or NT-pro-BNP differs in patients with and without HF.

Source: www.effectivehealthcare.ahrq.gov
Published Online: December 20, 2011

Due to the vast amount of literature published after January 2005, the obsolescence of certain assay types used in earlier studies of BNP and NT-pro-BNP test performance, as well as new or modified KQs that account for the evolution of (and continued uncertainty within) the field, an entirely new systematic review is required to provide a 'state of the art' assessment of this test in patients with HF. An update to the 2006 review, rather than a de novo review, would be a more limited contribution to the field.

To summarize the current body of scientific knowledge, this review will examine the diagnostic and prognostic use of BNP and NT-proBNP in several aspects of HF. The review will consider BNP and NT-proBNP test performance, cut points, and factors that affect test performance in emergency, urgent care, and primary care settings. As well, the review will investigate whether BNP and NT-proBNP are independent predictors of morbidity and mortality, or whether they add information to other methods used to predict morbidity and mortality. The review will examine whether therapies involving BNP and NT-proBNP improve outcomes in HF patients and whether the biologic variation of BNP and NT-proBNP differs in HF and non-HF populations.

In the following sections, we expand on the scientific rationale for performing a systematic review on BNP and NT-proBNP in HF diagnosis, prognosis, and therapy.

Diagnosis

Congestive HF is a common condition, especially among the elderly, and one of the most common reasons for admission to hospital. The diagnosis of HF remains a difficult clinical challenge. The diagnosis is based upon a constellation of symptoms (breathlessness, fatigue, ankle swelling), and signs (tachycardia, tachypnea, rales, increased jugular venous pressure, hepatomegaly, edema), supported by objective evidence of structural abnormality of the heart (shown by abnormalities in the echocardiogram or chest X-ray). The role of the natriuretic peptides (BNP and NT-proBNP) has been examined and reviewed by a previous systematic review of the evidence.[4] This and other reviews suggest that natriuretic peptides, due to the high sensitivity of the test, have value in ruling out the presence of HF. The low specificity of the test, however, limits its use as a rule-in tool.

Clinical guidelines including the 2009 update to the American College of Cardiology/American Heart Association (ACC/AHA) 2005 guideline for the diagnosis and management of HF in adults (section 3.1)[3] indicate that the measurement of natriuretic peptides may be a useful addition to the standard diagnostic tools for the evaluation of suspected HF. All, however, caution users to the poor specificity and the potential influences of confounders such as age, ethnicity, and comorbidities (including renal disease and obesity).

Since the publication of the AHRQ review in 2006, [4] several primary publications have addressed the diagnostic test accuracy of the natriuretic peptides for patients presenting to the emergency department and to primary care physicians.[5-11]

Re-evaluation of the use of the natriuretic peptides for the diagnosis of HF in light of new research findings is important to establish the strengths and weaknesses of this test (including

any adverse events). Both the emergent population (those with symptoms acute enough to warrant presentation to the emergency department or urgent care facilities) and the primary care population (those with risk factors, signs, and symptoms evaluated by a primary care physician) are areas of research that would benefit from a systematic review of the evidence. Decision cut-points have been proposed in several publications (most recently in the NICE Clinical Guideline No 108, 2010[12]), however they have not been optimized for specific populations and the effect of co-morbidities on the decision cut-points have not been systematically reviewed in terms of diagnosis. Examination of the optimal decision cut-points to maximize the diagnostic criterion of interest (sensitivity, specificity, best combination, specific population) as well as the effect of comorbidities in patients being evaluated will help to further refine the value of these tests.

Prognosis

Prognostic use of BNP and NT-proBNP has been studied in a number of primary studies and has been the subject of at least four systematic reviews.[13-16] The most recent of these systematic reviews includes primary studies up to July 2009.[13] Although these systematic reviews differed in the eligible studies evaluated, there was consistent evidence that BNP and NT-proBNP were independent predictors of mortality and other cardiac outcomes in patients with HF. In addition, they suggest that a discharge or post treatment BNP and NT-proBNP is a better predictor of prognosis. There is also suggestion that they add useful information to the standard cardiovascular disease (CVD) risk assessment in certain populations. A recently published Health Technology Assessment[17] did not ask prognostic questions and only reviewed papers up to July 2006. The updated NICE guideline[12] for chronic HF mentions that higher BNP and NT-proBNP levels are associated with poor prognosis in HF (section 3.1.3). There is a recommendation for high priority research in the area of determining prognostic stratification (page 208) and lists outcomes to be observed.

Two systematic reviews published in 2006 and 2005[4,14] have evaluated the evidence that BNP and NT-proBNP are predictive of mortality and other cardiac events in patients with HF. Doust et al.[14] evaluated studies in patients with HF but also in those with no overt disease. Based on this review BNP was shown to be consistently associated with an increased relative risk (RR) of death, including for asymptomatic subjects. The second systematic review[4] employed broader eligibility criteria and included almost double the number of eligible studies. This review showed similar results to the Doust et al., review, indicating that baseline BNP or NT-proBNP levels were independent predictors of mortality across various cut points. A small number of studies comparing baseline and predischarge BNP levels were also identified and suggested differences in the predictive strength for the outcome of mortality were shown in this second systematic review.[4]

Subsequent to these systematic reviews, there are a number of new research papers that consider BNP or NT-proBNP as a prognostic marker. Some of these studies are in the general population (or unselected) and others in selected populations, like HF. The prognostic use of BNP or NT-proBNP could become clearer if the full extent of the literature is considered.

Therapy

Optimization of therapy for patients with HF remains challenging despite use of proven medications and/or devices. This is in part due to the difficulty in perceiving signs and symptoms associated with HF unless they are overt. Current practice guidelines are based on target doses used in clinical trials, but are not individualized for patients. Appropriate drug titration algorithms may also depend on factors such as age, disease severity, and other comorbidities. In addition, measurement of BNP or NT-proBNP has been advocated as biomarkers to guide treatment. This is because they are independently associated with prognosis[15] and decrease with effective therapy.[18] The question of whether biomarker assisted therapy (to achieve a concentration below a target value) or intensified therapy (adjustment of therapy based on a change in biomarker concentration) reduces mortality, rehospitalization, or quality of life, compared to usual care is unclear.

At the time the AHRQ report on BNP was produced, the large interventional trials to address this question had just begun, so minimal data were available. Since then nine randomized controlled trials (RCTs) have been completed and several more RCTs are currently underway. In the most recent systematic review,[19] eight RCTs were reviewed and concluded that BNP-guided therapy was beneficial and the RR of all-cause mortality was lower in this group compared to the usual care group (RR, 0.76; 95% CI, 0.63 to 0.91; p = 0.03). Since its publication, a few limitations of this meta-analysis have been noted.[20] The main issues were the absence of information from the included studies and a discussion which did not thoroughly explain the findings.

Furthermore, knowledge of the variation of a test measure is important when treatment is based on a difference between serial measurements. We do not currently know how much of a difference in BNP or NT-proBNP concentrations is clinically important. Variation in a test measure is a function of the analytical variation of the assay method (bias and precision) and the inherent biological variation of the molecule tested. The biological variation may also be a function of disease severity, gender, medications and comorbidity.

Several studies have collected data in an attempt to understand the magnitude of the variation of BNP and NT-proBNP.[21-23] These studies have looked at the within day, day to day, and week to week variation of BNP and NT-proBNP in healthy individuals and in patients with stable chronic HF. The biological variation for individuals (CV_I) was found to increase with time between measurements for both BNP and NT-proBNP. However, there is inconsistency between studies, method type, and statistical analysis method. The CV_I values reported were as low as 8 percent and up to 50 percent.

Summary

The purpose of this review is to examine the evidence for the use of BNP and NT-proBNP in the diagnosis of HF, prognosis of HF in patients or general populations, and BNP or NT-proBNP guided therapy in patients with HF. The review is intended to inform physicians, clinical chemists, and policy makers in the use of these tests.

Source: www.effectivehealthcare.ahrq.gov
Published Online: December 20, 2011

II. The Key Questions

Key Question 1 (KQ1): In patients presenting to the emergency department or urgent care facilities with signs or symptoms suggestive of HF:

a) What is the test performance of BNP and NT-proBNP for HF?
b) What are the optimal decision cut points for BNP and NT-proBNP to diagnose and exclude HF?
c) What determinants affect the test performance of BNP and NT-proBNP (e.g., age, gender, comorbidity)?

Key Question 2 (KQ2): In patients presenting to a primary care physician with risk factors, signs, or symptoms suggestive of HF:

a) What is the test performance of BNP and NT-proBNP for HF?
b) What are the optimal decision cut points for BNP and NT-proBNP to diagnose and exclude HF?
c) What determinants affect the test performance of BNP and NT-proBNP (e.g., age, gender, comorbidity)?

Key Question 3 (KQ3): In HF populations, is BNP or NT-pro BNP measured at admission, discharge or change between admission and discharge an independent predictor of morbidity and mortality outcomes?

Key Question 4 (KQ4): In HF populations, does BNP measured at admission, discharge or change between admission and discharge add predictive information to other prognostic methods?

Key Question 5 (KQ5): Is BNP or NT-pro BNP measured in the community setting an independent predictor of morbidity and mortality outcomes in general populations?

Key Question 6 (KQ6): In patients with HF, does BNP assisted therapy or intensified therapy compared to usual care, improve outcomes?

Key Question 7 (KQ7): What is the biological variation of BNP and NT-proBNP in patients with HF and without HF?

Public Comments:

The KQs were posted for public comment on the Effective Health Care Program website between June 25 and July 26, 2011. Three comments were received and all cited or summarized literature related to the current review. No comments addressed the design of the KQs or the scope of the review and thus, no changes were made to the KQs during protocol development.

PICOTS and eligibility criteria

Population

Inclusion for all KQ's: Adults (>18 years of age)

KQ1 Inclusion: patients presenting to Emergency Department (ED) or Urgent Care settings with signs or symptoms consistent with HF (HF)**. Exclusion:** We will exclude studies where all subjects are ≤18 years of age, subjects that arrive at the ED or urgent care area with already diagnosed acute HF or known exacerbation of stable chronic HF, and studies that include only subjects with specific conditions that may impact BNP results such as heart transplantation, obesity, hypertrophic cardiomyopathy, or valvular lesions.

KQ2 Inclusion: patients presenting to a Primary Care physician with signs or symptoms consistent with HF. (Primary care will be defined according to the American Academy of Family Physicians' definition[24]) **Exclusion:** We will exclude studies where all subjects are ≤18 years of age, subjects with known acute HF or known exacerbation of stable chronic HF, and studies that include only subjects with specific conditions that may impact BNP results such as heart transplantation, obesity, hypertrophic cardiomyopathy, or valvular lesions.

KQ3, KQ4 Inclusion: patients with all types of HF (with or without any co-morbidity). We will attempt to categorize the type of HF at data extraction (e.g., acute, chronic, chronic with acute exacerbation). **Exclusion:** Adults at risk of coronary artery disease (CAD), with CAD, other adults at risk of HF without documented HF (diabetes, renal failure etc.).

KQ5 Inclusion: Adults in a community setting with no disease specified for the study (a non-selected or general population). **Exclusion:** Any study where a specific disease has been used to include or exclude subjects (e.g., acute coronary syndrome (ACS), CAD, Diabetes, Renal Failure, etc.)/

KQ6 Inclusion: Patients being treated for chronic HF. **Exclusion**: admitted patients with known HF or patients with acute HF.

KQ7 Inclusion: Adults with and without HF.

Interventions and Prognostic Factors

KQ1 and KQ2: Food and Drug Administration (FDA) approved assay BNP or NT-proBNP at admission or discharge or change in BNP /NT-proBNP between admission and discharge. No restriction on the BNP or NT-proBNP decision cutpoint. **Exclusion:** Use of non-FDA approved assay or not a BNP or NT-proBNP assay (i.e., pre-proBNP or atrial natriuretic peptide (ANP) and the versions of ANP)**.** No restriction on the cut-point.

KQ3 and 4: BNP or NT-pro BNP measured at admission, discharge, or change between admission and discharge. No restriction on the cut-point.

KQ5: No restriction on the decision cut-point. **Exclusion:** Non-BNP or NT-proBNP assay .

KQ6: Medical therapy based on BNP or NT-proBNP concentration.

KQ7: Multiple measurements of BNP or NT-proBNP per subject.

Comparators

KQ1 to 2: Any method of diagnosing HF that does not use BNP or NT-proBNP. No gold standard diagnostic criteria exist in HF, so we will calculate sensitivity and specificity of BNP or NT-proBNP using whatever comparator methods are used in the included studies.

KQ3, KQ4: New York Heart Association **(**NYHA) functional classification of stages of HF[25], ejection fraction, degree of hyponatremia, decreasing peak exercise oxygen uptake, decreasing hematocrit, widened QRS interval on 12-lead electrocardiogram, chronic hypotension, resting tachycardia, renal insufficiency, intolerance to conventional therapy, and refractory volume overload[3] or risk prediction scores (e.g., Seattle HF Model[26]). **Exclusion:** No comparator exclusion. No prediction score exclusion.

KQ5: Any predictive scoring system (e.g., Framingham[27]). **Exclusion:** No comparator exclusion. No prediction score exclusion.

KQ6: Medical therapy based on usual care for HF patients.

KQ7: No comparators.

Outcomes

KQ1: Test performance characteristics (sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratio, and area under the Receiver Operator Characteristic Curve).

The effect of various decision cut points on the test performance characteristics and the effect of various determinants (e.g., age, gender, co-morbidities) on the test performance characteristics. Adverse events associated with administration of the test, or being exposed to the results. These can be specific to patients or generalizable to the health care system. **Exclusion:** No restriction.

KQ2: Test performance characteristics (sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratio, and area under the Receiver Operator Characteristic Curve).

The effect of various decision cut points on the test performance characteristics and the effect of various determinants (e.g., age, gender, co-morbidities) on the test performance characteristics. Adverse events associated with administration of the test, or being exposed to the results. These can be specific to patients or generalizable to the health care system. **Exclusion:** No restriction.

KQ3 to KQ6: Mortality including all cause and HF; Morbidity including hospitalization (including HF, all cause, planned, unplanned); change in NYHA class; Quality of life. We will employ a broad definition of hospitalization, which includes any episode of HF that requires admission to a hospital bed beyond the emergency room for any length of time. This will include hospitalization for an initial diagnosis, readmission, stabilization, investigation, etc.

KQ7: Calculation of biological variation.

Timing or Follow-up

KQ1 to KQ7: No restriction.

Setting

KQ1: Emergency or Urgent Care Departments

KQ2: Primary Care Settings

KQ3: Admitted to hospital

KQ4: Outpatient Clinic/Ambulatory care, Hospital setting, Family practice

KQ5: Primary Care (Community or Family Practice or equivalent). **Exclusion:** Not primary care (Specialized out-patient clinics, Emergency room, admitted patients)

KQ6: No restriction

KQ7: No restriction

III. Analytic Framework

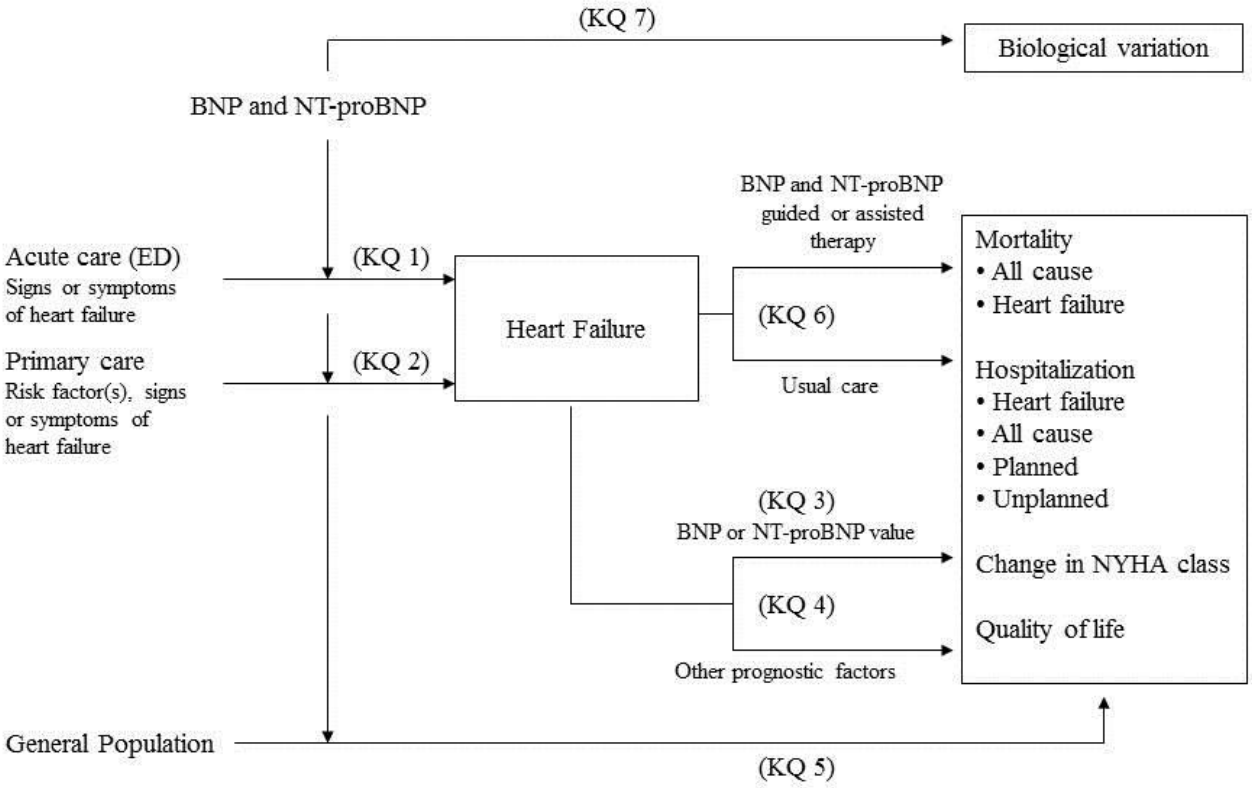

Figure 1. This figure is an analytic framework which depicts the KQs within the context of the PICOTS described in the previous section.

IV. Methods

A. Criteria for Inclusion/Exclusion of Studies in the Review

Inclusion and exclusion criteria will be based on the eligibility criteria from the PICOTS listed above in section two. All publications between 1989 and the current date will be included. Non-English publications will be excluded from this review. We will note the proportion of abstracts that are not in English. Study authors will be contacted via email for missing outcome or design data. Reference lists of included papers and any systematic reviews will be screened for potentially relevant papers that have not already been screened. Grey literature will be searched. See search strategy below in section B.

For KQ1, KQ2 and KQ7 the only excluded study design will be case reports. For KQ3 to KQ5, cross-sectional studies, case-control studies and retrospective studies

will be excluded. Only prospective designs will be included for KQ3 to KQ5 (RCT/CCT, cohort, before/after/ time series). For KQ6, only prospective randomized control trials RCT's will be included. For all KQ's, studies using non-FDA approved assays will be excluded. In addition letters, editorials, commentaries, abstracts from conference proceedings, will be excluded. Systematic reviews, meta-analyses will be excluded but reference lists will be evaluated for potentially relevant citations.

B. Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions

Search Strategy

The topic for this literature review is broad, looking at aspects of B-type natriuretic peptides and their use with HF diagnosis, monitoring, treatment and outcome. The search strategy (see Appendix A) will be based on that of an earlier review,[4] which was sufficiently broad for the current topic. Specifically, the search will use terms for b-type natriuretic peptides and will only be refined by date, language, and study subjects.

Search strategies will use combinations of controlled vocabulary (medical subject headings, keywords) and text words. The results will be captured from 1989 to the present (March 2011). The search will be restricted to human-focused studies (specifically removing those results that only include animal data), and to English language publications.
The search will be conducted in four databases: Medline, Embase, Cochrane Central and AMED. These databases have been chosen because they represent the best sources for a broad range of high quality literature relevant to this topic.

Review of reference lists of eligible studies at full text screening will be undertaken. Any potentially relevant citations will be cross-checked within our citation database. Any references not found within the database will be retrieved and screened at full text. We will not undertake hand searching for this review.

Grey Literature Search

The search strategy for grey literature will closely resemble the terms used in the bibliographic databases search. Terms for b-type natriuretic peptides will be searched with a focus on human studies. The aim of the grey literature search is to locate any information that has not yet reached more mainstream or indexed sources. As such, unpublished studies and ongoing studies are the focus of these searches.

Three types of grey literature sources will be searched: regulatory agency websites, clinical trial databases, and conference sources. The regulatory information includes the U.S Food and Drug Administration (FDA), Health Canada, and European Medicines Agency. The clinical trial databases that will be searched include: clinicaltrials.gov, clinicaltrialsregister.eu, metaRegister of Current Controlled Trials,

Clinical Trial Registries, Clinical Study Results, and WHO Clinical Trials. Conference papers will be searched in Conference Papers Index and Scopus for the last 2 years only. We will limit our conferences to the American Heart Association and the American College of Cardiology conferences.

Citations meeting our search criteria will be downloaded into Reference Manager Version 12 and then imported into systematic review software (DistillerSR 2011, Ottawa Ontario). Once in DistillerSR, citations will be screened using the specified eligibility criteria for the review.

Updating of the search
At the time of submission of the draft peer review report, an update of our search in all specified databases (see above) will be undertaken.

Incorporation of Public and Peer Review suggestions for literature
Any publications suggested by peer reviewers or from public comment will be documented and verified within our citation database. Any references not included within our citation database will be retrieved and screened for eligibility at full text.

C. Data Abstraction and Data Management

Relevant fields of information will be extracted from individual studies by trained data extractors using standardized forms and a reference guide. Prior to performing the data extraction, a calibration exercise will be conducted using a random sample of 10 included studies. All extraction will be reviewed by a study investigator and any disagreements will be resolved by consensus.

Extracted data for all studies will include study characteristics (e.g., first author, country of research origin, study design and sample size. Details of the patient population will include but not be limited to age, gender (% female), racial composition and co-morbidities. The following will also be extracted: blood sample type for BNP measurement (plasma or serum)**,** assay source (name)**,** type of peptide assessed (BNP, NTproBNP or both)**,** storage temperature of BNP (if applicable). For outcomes**,** we will extract the type of instrument or scale, cut points, primary or secondary outcome status, type of effect measure (endpoint or change score, measure of variance (standard deviation, standard error, etc.) and definition of treatment response.

For KQs 1 and 2 related to diagnosis, we will also extract the location of care (emergency/urgent care, primary care), information regarding the reference standard, test performance characteristics (either primary data to allow us to calculate these characteristics, or the summary data for sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratio, ROC curves) at various decision points and for various subgroups (age, gender, co-morbidities). Adverse events will be extracted if identified.

For the prognosis KQs 3, 4, and 5, data will be extracted for, HF score (NHYA or AHA/ACC), acute (and acute on chronic) or chronic HF, ejection fraction, other prognostic marker used as a comparator (degree of hyponatremia, decreasing peak exercise oxygen uptake, decreasing hematocrit, widened QRS on 12-lead electrocardiogram, chronic hypotension, resting tachycardia, renal insufficiency, intolerance to conventional therapy, and refractory volume overload), the study design (association with outcome; effect of BNP measurement on outcome; effect of BNP within a composite score on outcome), predefined confounders (age, NYHA, AHA/ACC, left ventricular ejection fraction (LVEF), degree of hyponatremia, decreasing peak exercise oxygen uptake, decreasing hematocrit, widened QRS interval on 12-lead electrocardiogram, chronic hypotension, resting tachycardia, and renal insufficiency), timing of BNP testing, BNP decision points used (cut-point), derivation of BNP cut-points, prevalence, length of follow-up, outcome (as per PICOTS), univariate (unadjusted) or multivariable (adjusted) analysis:- HR analysis with 95% CI; odds ratio(OR) with 95% CI; RR with 95% CI; Kaplan-Meier curves, Chi squared, likelihood ratios with 95% CI; detection probability; false referral probability and any other predictive scores or continuous outcome measurement. Variables used in models.

For KQ6, we will also extract a description of treatment arms (usual care, guided therapy, other); length of follow-up; blinding strategy; primary endpoint; secondary endpoint(s); HF etiology; percent of patients achieving target dose of medications in each study arm; statistical methods; adjustment factors; BNP or NT-proBNP concentrations at baseline and other time points, including change values; relative risk (RR), all groups reported.

For KQ7 we will extract the number of sequential measurements per subject; time between blood collections (e.g., hour, day, week, month, year);study length; sample collection parameters (e.g., tube type, handling, processing, storage);statistical methods to calculate coefficient of variation (CV), correlation, multivariate regression; CV, analytical (CVa);CV, individual (CVi);CV, between individual (CVg);Relative change value (RCV);Index of individuality (IOI) and factors associated with biological variation of BNP or NT-proBNP

Assessment of Methodological Quality of Individual Studies

To assess individual study quality, we will use methods recommended by the Agency for Healthcare Research and Quality's Evidence-based Practice Program in the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*[28] and the *Methods Guide for Medical Test Reviews.*[29] Two raters will assess the quality of individual studies using standardized quality assessment tools. We will minimize inconsistency amongst raters by providing standardized instructions and clear decision rules. Disagreements between raters will be resolved by consensus.

Quality assessment tools consist of five domains: population, outcome, exposure, statistical analysis and for RCT's, randomization, blinding and withdrawals. These

domains were adapted from the Newcastle-Ottawa quality assessment scales for case-control studies and cohort studies,[30] the Jadad scale[31]for RCT studies, QUADAS[32] for diagnostic tests, and Hayden et al[33] for prognosis studies. Additional items were needed to describe the population for case-control studies (five items), cohort studies (2 items), and before-after studies (2 items). Each quality item will be scored as yes, no or unsure. An answer of 'no' corresponds to a high risk of bias, 'unsure' corresponds to a possible or unclear risk of bias and 'yes' corresponds to a low risk of bias. For each quality item, we will graph the responses and discuss any problem areas. An overall quality score will not calculated. See Appendix B for a copy of the quality assessment tools.

E. Data Synthesis

Qualitative synthesis

Study results will be presented in four key sections based on diagnosis (KQ1 and 2), prognosis (KQ3 to 5) , treatment (KQ6) and biological variation (KQ7). All included studies will be summarized in narrative form and summary tables will be created showing key study characteristics (i.e. population characteristics, BNP test features, study outcomes,sample sizes, settings, funding sources, comparator treatments (type, dose, duration, and provider), methodological limitations, and any other important aspect related to each KQ. When studies cannot be meta-analyzed, graphical representation may be used to display main study outcomes.

Quantitative synthesis

The decision to pool individual study results will be based on clinical judgment with regards to comparability of study populations, diagnostic standard, treatments, and outcome measures. If meta-analysis is warranted, we will utilize the generic inverse variance method in Review Manager v5.1.2 (The Nordic Cochrane Centre, Copenhagen, Denmark) and the DerSimonian and Laird random and fixed effects models[34] to generate summary measures of effect (odds ratios) for each outcome. We will use Borenstein et al.'s formulae to convert data in other formats besides odds ratios (e.g., mean differences) into log odds ratios for input into Review Manager.[35] We will employ the I^2 test to assess statistical heterogeneity of summary odds ratios; results will be considered significant at the 5% level. Where homogeneity is present, we will report summary results with the fixed effects model. Otherwise, we will use the random effects model. Subgroup analyses for age, gender and co-morbidities will be undertaken to assess clinical diversity in the included studies.

F. Grading the Evidence for Each Key Question

We will assess the overall strength of the evidence (SOE) for KQ1, KQ2 and KQ6. The current system for SOE is not applicable for prognosis and biological variation studies. So for KQ3 to KQ5, that address prognosis, we will summarize the SOE

considering risk of bias, directness, and consistency. For KQ7, evaluating biological variation, SOE will not be evaluated.

For KQ1 and KQ2, we will assess SOE using the method for medical tests[29]and for KQ6 we will use the AHRQ method for intervention studies[28,36] The SOE will be classified into four grades based on the AHRQ approach: high, moderate, low, or insufficient.[28,36]

There are several domains of quality across studies that may influence the overall SOE for these KQs, including:

1) Risk of bias (how the study type and study design and conduct may have contributed to systematic error [bias])
2) Consistency of results (concerns homogeneity in direction and magnitude of results across different studies). In the context of the diagnostic test studies, this is typically considered as the spread of data points, for example on a ROC curve. In the context of intervention studies, this is the degree of spread of the summary effect size.
3) Directness of the evidence (concerns whether the evidence being assessed "reflects a single, direct link between the interventions of interest [medical tests] and the ultimate health outcome under consideration). In the context of diagnostic accuracy studies, there are unlikely to be any intermediate outcomes that would reduce the directness from the test being evaluated to the accuracy outcome. Directness also applies to comparing interventions and for both diagnostic accuracy and intervention studies, consideration should be given to how similar the test or the treatment is being used in practice.
4) Precision (refers to the width of confidence intervals for diagnostic accuracy outcomes, and the effect size for treatment monitoring; this domain is related to study sample size)
5) Other key domains (publication bias, dose-response association, existence of plausible unmeasured confounders, and strength of association [i.e., magnitude of effect]).

G. Assessing Applicability

Applicability may be affected by differences between what occurs in research and what happens in everyday clinical practice. We will assess applicability in accordance with the *Methods Guide for Medical Test Reviews.*[29]

The basis for applicability assessment in this review will be the populations studied and the settings of these populations. BNP and NT-proBNP values are known to differ between general populations and hospital based populations. These values are also correlated with increasing age and severity of disease.

The applicability of our findings will be limited to the populations and settings described in the protocol and PICOTS. We may encounter further limitations to applicability, depending on the actual populations and settings of the studies included in the review. We will address the applicability of these studies in terms of the degree to which the study populations and settings are relevant to individuals who would be considered for BNP or NT-pro-BNP testing in everyday clinical practice. We will also identify knowledge gaps related to population and setting (e.g., relevant subpopulations or settings that were not covered by our included studies).

V. References

1. Roger VL, Go AS, Lloyd-Jones DM, et al. Heart disease and stroke statistics--2011 update: a report from the American Heart Association. Circulation 2011 Feb 1;123(4):e18-e209
2. U.S. and Population Clocks. U.S.Census Bureau. 2011;
3. Jessup M. Focused update: ACCF/AHA guidelines for the diagnosis and management of heart failure in adults. J Am Coll Cardiol 2009;53(15):1343-82.
4. Balion C, Santaguida PL, Hill S, et al. Testing for BNP and NT-proBNP in the diagnosis and prognosis of heart failure. Evid Rep Technol Assess 2006;(142):1-147.
5. de Lemos JA, McGuire DK, Khera A, et al. Screening the population for left ventricular hypertrophy and left ventricular systolic dysfunction using natriuretic peptides: Results from the Dallas Heart Study. Am Heart J 2009;157(4):746-53.
6. Betti I, Castelli G, Barchielli A, et al. The role of N-terminal PRO-brain natriuretic peptide and echocardiography for screening asymptomatic left ventricular dysfunction in a population at high risk for heart failure. The PROBE-HF study. J Card Fail 2009;15(5):377-84.
7. Zuber M. Value of brain natriuretic peptides in primary care patients with clinical diagnosis of chronic heart failure. Scand Cardiovasc J 2009;43(5):324-9.
8. Rogers RK, Collins SP, Kontos MC, et al. Diagnosis and characterization of left ventricular hypertrophy by computerized acoustic cardiography, brain natriuretic peptide, and electrocardiography. J Electrocardiol 2008;41(6):518-25.
9. Behnes M, Brueckmann M, Ahmad-Nejad P, et al. Diagnostic performance and cost effectiveness of measurements of plasma N-terminal pro brain natriuretic peptide in patients presenting with acute dyspnea or peripheral edema. Int J Cardiol 2009;135(2):165-74.
10. Rutten JH, Steyerberg EW, Boomsma F, et al. N-terminal pro-brain natriuretic peptide testing in the emergency department: beneficial effects on hospitalization, costs, and outcome. Am Heart J 2008;156(1):71-7.
11. Kosowsky JM, Weiner C, Aronson AA, et al. Impact of point-of-care B-type natriuretic peptide (BNP) measurement on medical decision-making for older emergency department patients with dyspnea. J Emerg Med 2006;31(2):147-50.
12. National Clinical Guideline Centre. NICE Clinical Guideline 108: Chronic heart failure: Management of chronic heart failure in adults in primary and secondary care. Internet. http://guidance.nice.org.uk/CG108/Guidance/pdf/English

13. Di Angelantonio E., Chowdhury R, Sarwar N, et al. B-type natriuretic peptides and cardiovascular risk: systematic review and meta-analysis of 40 prospective studies. Circulation 2009;120(22):2177-87.

14. Doust JA, Pietrzak E, Dobson A, et al. How well does B-type natriuretic peptide predict death and cardiac events in patients with heart failure: systematic review. BMJ 2005;330(7492):625.

15. Vesbianu D, Vesbianu C, Bernstein P, et al. Plasma brain natriuretic peptide - An independent predictor of mortality and rehospitalization in congestive heart failure - A meta-analysis. World Heart J 2008;1(4):349-54.

16. Oremus M, Raina PS, Santaguida P, et al. A systematic review of BNP as a predictor of prognosis in persons with coronary artery disease. Clin Biochem 2008;41(4-5):260-5.

17. Mant J, Doust J, Roalfe A, et al. Systematic review and individual patient data meta-analysis of diagnosis of heart failure, with modelling of implications of different diagnostic strategies in primary care. Health Technol Asses 2009;13(32):1-207.

18. Balion CM, McKelvie RS, Reichert S, et al. Monitoring the response to pharmacologic therapy in patients with stable chronic heart failure: Is BNP or NT-proBNP a useful assessment tool? Clin Biochem 2008;41(4-5):266-76.

19. Porapakkham P, Porapakkham P, Zimmet H, et al. B-type natriuretic peptide-guided heart failure therapy: A meta-analysis. Arch Intern Med 2010;170(6):507-14.

20. Richards AM. B-type natriuretic peptide-guided therapy for chronic heart failure reduces all-cause mortality compared with usual care but does not affect all-cause hospitalisation or survival free of hospitalisation. Evid Based Med 2010;15(5):137-8.

21. Schou M, Gustafsson F, Nielsen PH, et al. Unexplained week-to-week variation in BNP and NT-proBNP is low in chronic heart failure patients during steady state. Eur J Heart Fail 2007;9(1):68-74.

22. Schou M, Gustafsson F, Kjaer A, et al. Long-term clinical variation of NT-proBNP in stable chronic heart failure patients. Eur Heart J 2007;28(2):177-82.

23. Wu AH. Serial testing of B-type natriuretic peptide and NTpro-BNP for monitoring therapy of heart failure: The role of biologic variation in the interpretation of results. Am Heart J 2006;152(5):828-34.

24. American Academy of Family Physicians. Primary Care. Internet. http://www.aafp.org/online/en/home/policy/policies/p/primarycare.html#Parsys0002

25. The Criteria Committee of the New York Heart Association. Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels. 9. Boston: Little, Brown and Co.; 1994.

26. Levy WC, Mozaffarian D, Linker DT, et al. The Seattle Heart Failure Model: Prediction of survival in heart failure. Circulation 2006;113(11):1424-33.

27. Kannel WB, D'Agostino RB, Silbershatz H, et al. Profile for estimating risk of heart failure. Arch Intern Med 1999;159(11):1197-204.

28. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality; 2011.

29. Agency for Healthcare Research and Quality. Methods Guide for Medical Test Reviews. Rockville, MD: Agency for Healthcare Research and Quality; 2010.

30. Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., and Tugwell, P. The Newcastle-Ottawa Scale (NOW) for assessing the quality of nonrandomised studies in meta-analyses. Internet. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

31. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Control Clin Trials 1996;17(1):1-12.

32. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003;3:25.

33. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med 2006;144(6):427-37.

34. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.

35. Borenstein M, Hedges LV, Higgins JP, et al. Introduction to Meta-Analysis. United Kingdom: John Wiley & Sons Ltd; 2009.

36. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: Grading the strength of a body of evidence when comparing medical interventions--Agency for Healthcare Research and Quality and the Effective Health-care program. J Clin Epidemiol 2010;63(5):513-23.

VI. Definition of Terms

No additional definitions to those provided in the text above.

VII. Summary of Protocol Amendments

In the event of protocol amendments, the date of each amendment will be accompanied by a description of the change and the rationale.

VIII. Review of Key Questions

For all EPC reviews, KQs were reviewed and refined as needed by the EPC with input from Key Informants and the Technical Expert Panel (TEP) to assure that the questions are specific and explicit about what information is being reviewed. In addition, for Comparative Effectiveness reviews, the KQs were posted for public comment and finalized by the EPC after review of the comments.

IX. Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the KQs for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The Task Order Officer (TOO) and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts comprise a multi-disciplinary group of clinical, content, and methodologic experts who provide input in defining populations, interventions, comparisons, or outcomes as well as identifying particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicted opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design and/or methodological approaches do not necessarily represent the views of individual technical and

content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor contribute to the writing of the report and have not reviewed the report, except as given the opportunity to do so through the public review mechanism

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. Peer review comments on the preliminary draft of the report are considered by the EPC in preparation of the final draft of the report. Peer reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will, for CERs and Technical briefs, be published three months after the publication of the Evidence report.

Potential Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than $10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

Appendix A
Literature Search Strategy

Medline
Database: Ovid MEDLINE(R) <1948 to February week 4 2011>, Ovid MEDLINE(R) Daily Update <March 08, 2011>, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <March 08, 2011>
Search Strategy:
---

1 natriuretic peptide, brain/
2 bnp.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]
3 nt-probnp.mp.
4 brain-type natriuretic peptide.mp.
5 bnp1-32.mp.
6 bnp-32.mp.
7 bnp77-108.mp.
8 probnp.mp.
9 nt-probnp1-76.mp.
10 natriuretic factor-32.mp.
11 natriuretic peptide type-b.mp.
12 type-b natriuretic peptide.mp.
13 ventricular natriuretic peptide.mp.
14 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13
15 14
16 limit 14 to ed=19890101-20110309
17 human/
18 animal/
19 17 and 18
20 18 not 19
21 16 not 20
22 16 and 17
23 21 not 22
24 limit 23 to english language

EMBASE
Database: EMBASE <1988 to 2011 Week 09>
Search Strategy:
---

1 Brain Natriuretic Peptide/ct, ec, an, dv [Clinical Trial, Endogenous Compound, Drug Analysis, Drug Development]
2 bnp.tw.
3 nt-probnp.tw.
4 brain-type natriuretic peptide.tw.

5 bnp 1-32.tw.
6 bnp1-32.tw.
7 bnp-32.tw.
8 bnp77-108.tw.
9 bnp 77-108.tw.
10 probnp.tw.
11 nt-probnp1-76.tw.
12 nt-probnp 1-76.tw.
13 natriuretic factor-32.tw.
14 natriuretic peptide type-b.tw.
15 type-b natriuretic peptide.tw.
16 ventricular natriuretic peptide.tw.
17 or/1-16
18 ("1989$" or "1990$" or "1991$" or "1992$" or "1993$" or "1994$" or "1995$" or "1996$" or "1997$" or "1998$" or "1999$" or "2000$" or "2001$" or "2002$" or "2003$" or "2004$" or "2005$" or "2006$" or "2007$" or "2008$" or "2009$" or "2010$" or "2011$").ew.
19 17 and 18
20 limit 19 to (human and english language)
21 limit 20 to "review"
22 20 not 21
23 Brain Natriuretic Peptide/ct, ec, an, dv [Clinical Trial, Endogenous Compound, Drug Analysis, Drug Development]
24 bnp.tw.
25 nt-probnp.tw.
26 brain-type natriuretic peptide.tw.
27 bnp 1-32.tw.
28 bnp1-32.tw.
29 bnp-32.tw.
30 bnp77-108.tw.
31 bnp 77-108.tw.
32 probnp.tw.
33 nt-probnp1-76.tw.
34 nt-probnp 1-76.tw.
35 natriuretic factor-32.tw.
36 natriuretic peptide type-b.tw.
37 type-b natriuretic peptide.tw.
38 ventricular natriuretic peptide.tw.
39 or/23-38
40 ("1989$" or "1990$" or "1991$" or "1992$" or "1993$" or "1994$" or "1995$" or "1996$" or "1997$" or "1998$" or "1999$" or "2000$" or "2001$" or "2002$" or "2003$" or "2004$" or "2005$" or "2006$" or "2007$" or "2008$" or "2009$" or "2010$" or "2011$").ew.
41 39 and 40
42 human/
43 animal/

44 animal experiment/
45 43 or 44
46 42 and 45
47 45 not 46
48 41 and 47
49 41 not 48
50 49 not 22
51 limit 50 to english language

AMED
Database: AMED (Allied and Complementary Medicine) <1985 to February 2011>
Search Strategy:
---
1 exp peptides/
2 bnp.tw.
3 nt-probnp.tw.
4 brain-type natriuretic peptide.tw.
5 bnp 1-32.tw.
6 bnp-32.tw.
7 bnp77-108.tw.
8 probnp.tw.
9 nt-probnp1-76.tw.
10 natriuretic factor-32.tw.
11 natriuretic peptide type-b.tw.
12 type-b natriuretic peptide.tw.
13 ventricular natriuretic peptide.tw.
14 or/1-13
15 limit 14 to yr=1989-Current

CCRT
Database: EBM Reviews - Cochrane Central Register of Controlled Trials <1st Quarter 2011>
Search Strategy:
---
1 Natriuretic Peptide, Brain/me, bi, bl, se, du [Metabolism, Biosynthesis, Blood, Secretion, Diagnostic Use]
2 bnp.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
3 nt-probnp.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
4 brain-type natriuretic peptide.tw.
5 bnp1-32.tw.
6 bnp-32.tw.
7 bnp77-108.tw.
8 probnp.tw.
9 nt-probnp1-76.tw.
10 natriuretic factor-32.tw.

11 natriuretic peptide type-b.tw.
12 type-b natriuretic peptide.tw.
13 ventricular natriuretic peptide.tw.
14 or/1-13
15 limit 14 to yr=1989-Current

CDSR
Database: EBM Reviews - Cochrane Database of Systematic Reviews <1989 to March 2011>
Search Strategy:
---
1 [Natriuretic Peptide, Brain/me, bi, bl, se, du [Metabolism, Biosynthesis, Blood, Secretion, Diagnostic Use]]
2 bnp.mp. [mp=title, abstract, full text, keywords, caption text]
3 nt-probnp.mp. [mp=title, abstract, full text, keywords, caption text]
4 brain-type natriuretic peptide.tw.
5 bnp1-32.tw.
6 bnp-32.tw.
7 bnp77-108.tw.
8 probnp.tw.
9 nt-probnp1-76.tw.
10 natriuretic factor-32.tw.
11 natriuretic peptide type-b.tw.
12 type-b natriuretic peptide.tw.
13 ventricular natriuretic peptide.tw.
14 or/1-13
15 limit 14 to yr=1989-Current

CINAHL
Database: CINAHL - Cumulative Index to Nursing & Allied Health Literature <1982 to March Week 1 2011>
Search Strategy:
---
1 exp Peptides/an, me, bl, ph, st, df, du, ur [Analysis, Metabolism, Blood, Physiology, Standards, Deficiency, Diagnostic Use, Urine]
2 nt-probnp.tw.
3 brain-type natriuretic peptide.tw.
4 bnp 1-32.tw.
5 bnp-32.tw.
6 bnp77-108.tw.
7 probnp.tw.
8 nt-probnp1-76.tw.
9 natriuretic factor-32.tw.
10 natriuretic peptide type-b.tw.
11 type-b natriuretic peptide.tw.

12 ventricular natriuretic peptide.tw.
13 or/1-12
14 limit 13 to yr=1989-2011

Appendix B
Quality Assessment Forms

Quality Assessment Form: Case Control Design

Study Population	**Yes/No/Unsure**
1. Was the definition of the case group adequate, i.e., inclusion and/or exclusion criteria described (no specific criteria)?	
2. Were the participants in the case group representative of the case definition?	
3. Was the definition of the control group adequate, i.e., inclusion and/or exclusion criteria described (no specific criteria)?	
4. Were the participants in the control group representative of the control definition?	
Comparability	
5. Were the study participants in the case and control groups recruited over the same time period?	
Outcome Measurements	
6. Was the outcome defined clearly (i.e., was the measure described in sufficient detail to be replicated)?	
7. Were those measuring the main outcome blind to exposure status?	
Exposure Measurements	
8. Was the exposure defined clearly (i.e., was the test method described in sufficient detail to permit replication)?	
9. Were those measuring the exposure blind to outcome status?	
Statistical analysis	
10. Were potential confounders (i.e., age, gender) measured and adequately addressed in the analysis?	
11. Was the statistical analysis described?	
12. Were missing data reported?	

Quality Assessment Form: Cohort Design

Study Population	**Yes/No/Unsure**
1. Did the authors clearly describe the population from which the participants were drawn?	
2. Were the inclusion and/or exclusion criteria described (no specific criteria)?	
3. Were the participants in the study representative of the population from which they were recruited?	
Outcome Measurements	
4. Was the outcome defined clearly (i.e., was the measure described in sufficient detail to be replicated)?	
5. Were those measuring the main outcome blind to exposure status?	
6. Was the outcome absent at the beginning of the study?	
7. Was the length of the study long enough for the outcome to occur?	
8. Was the follow-up adequate, i.e., were losses documented?	
Exposure Measurements	
9. Was the exposure defined clearly (i.e., was the test method described in sufficient detail to permit replication)?	
10. Were those measuring the exposure blind to outcome status?	
Statistical analysis	
11. Were potential confounders (i.e., age, gender) measured and adequately addressed in the analysis?	
12. Was the statistical analysis described?	
13. Were missing data reported?	

Quality Assessment Form: Randomized Control Design

Study Population	**Yes/No/Unsure**
1. Did the authors clearly describe the population from which the participants were drawn?	
2. Were the inclusion and/or exclusion criteria described (no specific criteria)?	
3. Were the participants in the study representative of the population from which they were recruited?	

Randomization	
4. Was randomization appropriate?	
5. Was double blinding reported?	
6. Was double blinding appropriate?	
7. Were withdrawals reported by number and reason per arm?	
Outcome Measurements	
8. Was the outcome defined clearly (i.e., was the measure described in sufficient detail to be replicated)?	
9. Were those measuring the main outcome blind to exposure status?	
Exposure Measurements	
10. Was the exposure defined clearly (i.e., was the test method described in sufficient detail to permit replication)?	
11. Were those measuring the exposure blind to outcome status?	
Statistical analysis	
12. Were potential confounders measured and adequately addressed in the analysis?	
13. Was the statistical analysis described?	
14. Were missing data reported?	

Quality Assessment Form: Cross-Sectional Design and Before-After Design

Study Population	**Yes/No/Unsure**
1. Did the authors clearly describe the population from which the participants were drawn?	
2. Were the inclusion and/or exclusion criteria described (no specific criteria)?	
3. Were the participants in the study representative of the population from which they were recruited?	
Outcome Measurements	
4. Was the outcome defined clearly (i.e., was the measure described in sufficient detail to be replicated)?	

Source: www.effectivehealthcare.ahrq.gov
Published Online: December 20, 2011

5. Were those measuring the main outcome blind to exposure status?	
Exposure Measurements	
6. Was the exposure defined clearly (i.e., was the test method described in sufficient detail to permit replication)?	
7. Were those measuring the exposure blind to outcome status?	
Statistical analysis	
8. Were potential confounders measured and adequately addressed in the analysis?	
9. Was the statistical analysis described?	
10. Were missing data reported?	
Before-After Design	
Comparability	
1. Did the authors ensure comparability of groups by addressing selection from same source population, (i.e., both groups from the same setting)?	
2. Did the authors ensure comparability of groups by reporting participation rates by group?	
3. Were the study participants in different comparison groups recruited over the same time period?	

Quality Assessment for Prognosis Papers: Hayden Criteria

Study Participation	**Yes/No/Unsure**
1. The study sample represents the population of interest on key characteristics, sufficient to limit potential bias of the results.	
Study Attrition	
2. Loss to follow-up (from sample to study population) is not associated with key characteristics (i.e., the study data adequately represent the sample), sufficient to limit potential bias.	
Prognostic Factor Measurement	
3. The prognostic factor of interest is adequately measured in study participants to sufficiently limit potential bias	
Outcome Measurement	
4. The outcome of interest is adequately measured in study participants to sufficiently limit potential bias	
Confounding measurement and account	

Source: www.effectivehealthcare.ahrq.gov
Published Online: December 20, 2011

5. Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest.	
Analysis	
6. The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid results.	